Matrix Exponentiation

Miliyon T.

August 25, 2016

1 Introductions

The matrix exponential plays an important role in solving system of linear differential equations. On this page, we will define such an object and show its most important properties. The natural way of defining the exponential of a matrix is to go back to the exponential function e^x and find a definition which is easy to extend to matrices. Indeed, we know that the Taylor polynomials

$$T_n(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$$

converges pointwise to e^x and uniformly whenever x is bounded. These algebraic polynomials may help us in defining the exponential of a matrix. Indeed, consider a square matrix A and define the sequence of matrices

$$A_n = I_n + \frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots + \frac{1}{n!}A^n.$$

When n gets large, this sequence of matrices get closer and closer to a certain matrix. This is not easy to show; it relies on the conclusion on e^x above. We write this limit matrix as e^A. This notation is natural due to the properties of this matrix. Thus we have the formula

$$e^A = I_n + \frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots + \frac{1}{n!}A^n + \cdots\cdots.$$

One may also write this in series notation as

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!}A^n$$

At this point, the reader may feel a little lost about the definition above. To make this stuff clearer, let us discuss an easy case: diagonal matrices.

2 Examples

Example 2.1. Consider the diagonal matrix

$$A = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}.$$

It is easy to check that

$$A^n = \begin{pmatrix} 2^n & 0 \\ 0 & (-1)^n \end{pmatrix}$$

for $n = 1, 2, \cdots$. Hence we have

$$I_n + \frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots + \frac{1}{n!}A^n = \begin{pmatrix} 1 + \frac{2}{1!} + \cdots + \frac{2^n}{n!} & 0 \\ 0 & 1 + \frac{(-1)}{1!} + \cdots + \frac{(-1)^n}{n!} \end{pmatrix}.$$

Using the above properties of the exponential function, we deduce that

$$e^A = \begin{pmatrix} e^2 & 0 \\ 0 & e^{-1} \end{pmatrix}.$$

Indeed, for a diagonal matrix A, e^A can always be obtained by replacing the entries of A (on the diagonal) by their exponentials. Now let B be a matrix similar to A. As explained before, then there exists an invertible matrix P such that

$$B = P^{-1}AP.$$

Moreover, we have

$$B^n = P^{-1}A^n P$$

for $n = 1, 2, \cdots$, which implies

$$I_n + \frac{1}{1!}B + \frac{1}{2!}B^2 + \cdots + \frac{1}{n!}B^n \ldots \ldots \frac{1}{1!}A + \frac{1}{2!}A^2 + \cdots + \frac{1}{n!}A^n \Big) P.$$

This clearly implies that

$$e^B = P^{-1} \begin{pmatrix} e^2 & 0 \\ 0 & e^{-1} \end{pmatrix} P.$$

In fact, we have a more general conclusion. Indeed, let A and B be two square matrices. Assume that $A \sim B$. Then we have $e^A \sim e^B$. Moreover, if $B = P^{-1}AP$, then

$$e^B = P^{-1}e^A P.$$

Example 2.2. Consider the matrix

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

This matrix is upper-triangular. Note that all the entries on the diagonal are 0. These types of matrices have a nice property. Let us discuss this for this example. First, note that

$$A^2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \mathcal{O}.$$

In this case, we have

$$e^A = I + A + \frac{1}{2!}A^2 = \begin{pmatrix} 1 & 1 & 3/2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

In general, let A be a square upper-triangular matrix of order n. Assume that all its entries on the diagonal are equal to 0. Then we have

$$A^n = \mathcal{O}.$$

Such matrix is called a nilpotent matrix. In this case, we have

$$e^A = I_n + \frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots + \frac{1}{(n-1)!}A^{n-1}.$$

As we said before, the reasons for using the exponential notation for matrices reside in the following properties:

Theorem 2.3. The following properties hold:

1. $e^{\mathcal{O}} = I_n$;

2. If A and B commute, meaning $AB = BA$, then we have

$$e^{A+B} = e^A e^B;$$

3. For any matrix A, e^A is invertible and

$$\left(e^A\right)^{-1} = e^{-A}.$$